3Ф

УДК 004.822: 004.823

П.І. Федорук, М.С. Дутчак
Прикарпатський національний університет імені Василя Стефаника
Україна, 76018, м. Івано-Франківськ, вул. Т.Г. Шевченка, 57

Інтелектуальний аналіз природномовних текстів у процесі формування бази знань адаптивних систем дистанційного навчання

P.I. Fedoruk, M.S.Dutchak
Carpathian National University named after V. Stefanik
Ukraine, 76018, Ivano-Frankivsk, Taras Shevchenko st., 57.

Intelligent Analysis of Natural-Language Texts in the Process of Building of Knowledge Base for Adaptive E-Learning

П.И. Федорук, М.С. Дутчак
Прикарпатский национальный университет имени Василя Стефаника
Украина, 76018, г. Ивано-Франковск, ул. Т.Г. Шевченко, 57

Интеллектуальный анализ естественно-языковых текстов в процессе формирования базы знаний адаптивных систем дистанционного обучения

У статті описано алгоритм і програмну реалізацію модуля автоматизованого квантування лекційного матеріалу на основі інтелектуального аналізу природномовних текстів в процесі побудови бази знань адаптивних систем дистанційного навчання. Результатом даного дослідження є спрощення процедури підготовки лекційного матеріалу до імпорту в БЗ; встановлення взаємозв’язків між окремими елементами БЗ, визначення властивостей і їх подальше використання для побудови індивідуальної траєкторії навчання.
Ключові слова: адаптивна система дистанційного навчання, морфологічний аналіз, база знань.

This article describes the algorithm and program realization of the module for automated quantization of the lecture material based on natural-language texts in the process of building of knowledge base for adaptive e-learning. The result of this research is simplification the preparation of lecture material to be imported into the knowledge base, establishing of relationships between the individual elements of the knowledge base, and determination of their properties for future use in the construction of an individual trajectory of study.
Key Words: adaptive system of distance learning, morphological analysis, knowledge base.

В статье описаны алгоритм и программная реализация модуля автоматизированного квантования лекционного материала на основе интеллектуального анализа естественно-языковых текстов в процессе построения базы знаний адаптивных систем дистанционного обучения. Результатом данного исследования является упрощение процедуры подготовки лекционного материала к импорту в базу знаний, установление взаимосвязей между отдельными элементами базы знаний и определение их свойств для дальнейшего использования при построении индивидуальной траектории обучения.
Ключевые слова: адаптивная система дистанционного обучения, морфологический анализ, база знаний.

Вступ

Особливістю сучасної системи освіти є можливість навчання за індивідуальною освітньою траєкторією, яку складно організувати з використанням тільки традиційних систем навчання. На допомогу традиційним системам приходять адаптивні системи дистанційного навчання (АСДН) [1]. Основою АСДН є база знань, від ефективної розробки і реалізації якої залежить ефективне функціонування всієї системи. Ба́за знань (БЗ) – це особливого роду база даних, розроблена для управління знаннями (метаданими), тобто збором, зберіганням, пошуком і видачею знань [2]. Виконання всіх цих етапів функціонування БЗ є особливо актуальним для АСДН. Більшість БЗ існуючих середовищ комп'ютерного навчання ґрунтуються на розміщенні в них електронних варіантів конспектів традиційних занять і статичній видачі навчального матеріалу незалежно від рівня знань і особливостей студента.

Огляд найбільш відомих систем, широко використовуваних в сучасному дистанційному навчанні (ДН), таких як ANGEL, BlackBoard, Desire2Learn, ILIAS, Lotus LearningSpace, Moodle, WebCT, TopClass, Zhang, OpenClass, EduPro, показав, що, як правило, навчальний курс, представлений в середовищі системи ДН, являє собою набір статичних гіпертекстових документів. Хоча можна виділити системи, в яких частково реалізовано принцип адаптивності. Наприклад, БЗ АСДН EduPro, яка розроблена Центром дистанційного навчання та контролю знань Прикарпатського національного університету імені Василя Стефаника, містить навчальний матеріал, структурований відповідно до власних стандартів і має можливість його адаптивної видачі відповідно до прогалин у знаннях студентів. Проте наповнення БЗ даної системи вимагає досить значних затрат часу і зусиль викладачів чи інженерів зі знань в процесі підготовки НМ для імпорту в БЗ відповідно до вимог стандарту EduPro [3], також в даній системі недостатньо досліджено і використано взаємозв'язок між окремими елементами БЗ.

Отже, більшість існуючих освітніх середовищ, платформ та порталів, використовуючи новітні комп'ютерні технології, зазвичай не є адаптивними. Для того щоб статичний гіпертекстовий документ став динамічним, в «інтелект» СДН необхідно закласти алгоритми розуміння і маніпулювання природномовними текстами, що є одним з найбільш відомих завдань в галузі штучного інтелекту.

Дослідження цієї проблеми почалося ще на рубежі 60-х – 70-х рр.; відтоді було зроблено безліч різних спроб її вирішення. Створені десятки експериментальних програм, здатних вести діалог з користувачем природною мовою.

Проте широкого розповсюдження такі системи досі не отримали – як правило, через невисоку якість розпізнавання фраз, жорсткі вимоги до синтаксису «природної мови», а також через великі витрати машинного часу і ресурсів, необхідних для їх роботи. Практично в усіх системах машинного розуміння тексту використовується обмежена природна мова, оскільки повної і строгої формальної моделі для жодної природної мови досі не створено.

Для формування БЗ АСДН важливим моментом є квантування лекційного матеріалу, яке полягає в його розбитті на складові частини, описі властивостей цих частин і встановленні зв'язків між ними. Тому **метою даного дослідження** є розробка ефективних алгоритмів інтелектуального аналізу природномовних текстів у процесі їх імпорту в БЗ АСДН.

Структурування лекційного матеріалу

В основу формування структури БЗ АСДН, її наповнення та визначення необхідних параметрів кожного з її елементів закладено фреймову модель подання знань.

Фреймова модель подання знань заснована на теорії фреймів М. Мінскі, відповідно до якої *фрейм* – це формалізована модель представлення образу [4], тобто він є описом деякого об'єкта. Фрейму присвоюється ім'я і ряд слотів, які містять атрибути або процедурні знання, пов'язані з його атрибутами. З кожним слотом може пов'язуватись умова на заповнення чи асоційовані процедури.

Більшість систем штучного інтелекту використовують набір фреймів, що з'єднані один з одним певним числом і утворюють певну ієрархію. Однією з найбільш важливих властивостей фреймів у таких ієрархіях є наслідування властивостей [2].

Для представлення знань у вигляді ієрархічної організації фреймів (ІОФ) весь навчальний матеріал розбивається на навчальні блоки (НБ) різних рівнів ієрархії:

1 рівень: навчальна дисципліна;

2 рівень: окремі заняття;

3 рівень: пункти плану заняття;

4 рівень: підпункти плану заняття або кванти навчальної інформації (КНІ);

5 рівень: підпункти плану заняття або КНІ;

…..

n-1 рівень: підпункти плану заняття або КНІ;

n рівень: КНІ.

Кількість рівнів залежить від глибини структурованості навчального матеріалу (НМ).

Для пунктів та підпунктів плану введено поняття незалежних навчальних блоків (ННБ), для них висувається вимога синтаксичної незв'язності з іншими НБ. В межах ННБ виділяють КНІ – сукупність тісно пов'язаної інформації, яка формулює певну думку, але при самостійній подачі може бути незрозумілою чи потребувати доповнення іншими КНІ ННБ.

Поділ на НБ має суб'єктивний характер, оскільки залежить від суб'єктивного рішення інженера зі знань чи будь-якого іншого суб'єкта, що здійснює структурування матеріалу.

Після поділу НМ на КНІ, між ними встановлюються наступні типи зв'язку: синтаксичний, змістовний та якісний. При традиційному навчанні всі ці зв'язки, як правило, закладені в «інтелекті» викладача чи частково викладені на папір. Для комп'ютеризованого навчання необхідна побудова «штучного інтелекту» освітніх систем з метою проведення максимально ефективного навчання з мінімальним втручанням у навчальний процес людини-викладача.

Для математичного вираження зв'язку між КНІ введено коефіцієнт залежності (КЗ) k. КЗ показує тісноту зв'язку між двома об'єктами за певними ознаками, він є відносною мірою і набуває значення від 0 до +1. Чим ближче значення k до 1, тим щільніший зв'язок. При $k = 0$ зв'язок відсутній.

Синтаксичний зв'язок – текст i-го КНІ побудовано так, що він синтаксично вимагає іншого j-го КНІ, без j-го КНІ i-й КНІ не може бути введений в НМ. Якщо j-й КНІ синтаксично залежить від i-го КНІ, то синтаксичний КЗ $sk_{j,i} = 1$, інакше $sk_{j,i} = 0$.

Змістовний зв'язок – КНІ пов'язані за змістом, але необов'язково за синтаксисом. У випадку, коли i-й КНІ синтаксично залежить від j-го КНІ, то він обов'язково залежить і змістовно. Значення змістових КЗ (zk) встановлюються з використанням інструментальних засобів систем одержання інформації з тексту, описаних нижче.

Якісний зв'язок – якість засвоєння одного КНІ залежить від якості засвоєння іншого КНІ. Значення якісного зв'язку виражається якісним КЗ (jk). Якісний КЗ доцільно визначати для змістовно зв'язних КНІ, оскільки, якщо j-й КНІ не пов'язаний змістовно з i-м КНІ безпосередньо чи опосередковано, i-й КНІ і не може мати впливу на якість засвоєння j-го КНІ [5].

Для проведення тесту на синтаксичну зв’язність КНІ необхідний словник синтаксично-зв’язних слів: опираючись на дану, альтернативою є, тому, при цьому, причому, таким чином, в обох випадках і.т.д., а також він, вона, воно, вони, його, їх і т.д., як вони зустрічаються в головному, а не у підрядному реченні. Якщо в межах *i*-го КНІ виявлено синтаксично-зв’язне слово, то *i*-й КНІ зв’язний синтаксично, як правило, з попереднім КНІ, тобто за яким він безпосередньо слідує.

Морфологічний аналізатор

Проведений морфологічний і синтаксичний аналіз текстів на природній мові показав, що, порівнюючи головні компоненти іменних груп, а при їх тотожності і цілі іменні групи НБ з іменними групами деякої послідовності НБ, можна одержати змістові КЗ між НБ. В результаті проведення морфологічного аналізу для кожного НБ визначаються кандидати на зіставлення. Ці кандидати являють собою іменники НБ, а також усі англомовні слова і абревіатури. Для визначення того, чи слово є іменником, розглянуто два способи. Перший спосіб – подати запит до морфологічного словника [6]. Другий спосіб – задання правил визначення приналежності текстової одиниці до тієї чи іншої частини мови. Другий спосіб потребує використання значно менших за обсягом словників, але вимагає написання достатньо великої кількості правил. Для того щоб застосування цих правил давало точний результат, потрібні не тільки значні затрати часу і зусиль програміста, але й залучення до роботи висококваліфікованого лінгвіста, що на практиці не завжди є виправданим. Причому затрати часу для пошуку в словнику в більшості випадків значно менші, ніж для проведення усіх необхідних тестів над текстовою одиницею. Тому для визначення слів-кандидатів на зіставлення використано перший спосіб.

Очевидно, для зберігання декількох десятків тисяч словникових записів зі складною структурою і взаємними посиланнями необхідно використати базу даних. Як СУБД для словникової підсистеми була вибрана PostgreSQL.

Крім того, використання СУБД для зберігання морфологічної інформації дозволяє звести завдання морфологічного аналізу до стандартного завдання пошуку у базі даних.

Загальна структура словникової бази даних зображена на рис. 1.

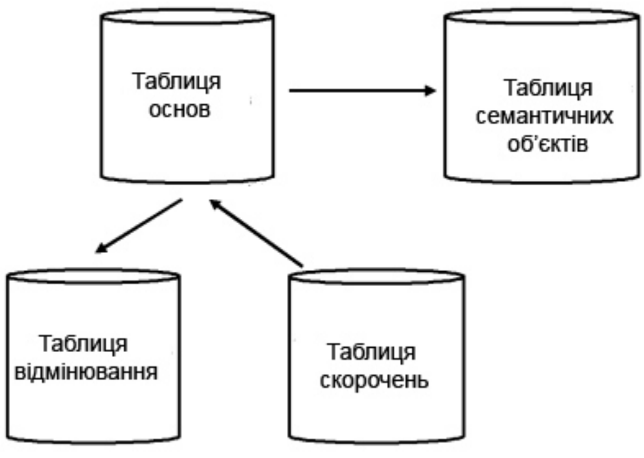


Рисунок 1 – Структура словникової бази даних

Кожна таблиця БД містить записи, що відповідають її назві. Крім цього, у таблиці відмінювання кожен запис ідентифікується частиною мови.

Для забезпечення більшої гнучкості при зберіганні семантичної інформації, пов’язаної з кожним словом, вводиться таблиця семантичних об’єктів. В результаті один

запис в таблиці основ природним чином представляє відразу усі омоніми цього слова – за рахунок декількох посилань на різні семантичні об'єкти. І навпаки, декілька синонімічних основ можуть містити посилання на один і той же семантичний об'єкт.

В словникову підсистему включена також таблиця скорочень, оскільки навіть в офіційних документах існує досить широкий набір загальноприйнятих скорочень, їх обробка виявляється необхідною. Запис в таблиці скорочень має посилання на запис в таблиці основ, яка містить повну форму цього слова.

Серед знайдених слів-кандидатів виключаються загальновживані слова, такі, наприклад, як «випадок», «приклад», «опис», «система», «формат» і т.д.

Наступним етапом є встановлення ЗКЗ між КНІ.

Тести для встановлення змістовно-зв'язних КНІ:

1. Якщо зв'язний синтаксично, то обов'язково і змістовно.

2. Застосування морфологічного аналізатора.

В основу розробленого морфологічного аналізатора закладено стандартну роботу із запитами до бази даних, що містить морфологічну інформацію і відомий метод дерева прийняття рішень (ДПР).

У загальному випадку дерево прийняття рішень – це неорієнтований зв'язний граф без циклів, вузлами якого є гіпотези, що перевіряються, а листям – можливі рішення. У міру просування від кореня дерева до листя відбувається послідовне «уточнення» рішення, що приймається [7]. Існує три ключові поняття, що стосуються ДПР:

– Передісторія – весь набір вхідної інформації, на підставі якої потрібно прийняти рішення. Стосовно морфологічного аналізу передісторія включає в себе поточний стан дерева морфологічного розбору, в тому числі морфологічну інформацію про його вершини, їх взаємне положення в тексті і т.д.

– Набір питань (і можливих відповідей на них) – евристичні знання про те, що в передісторії є істотним для прийняття рішення. Кожен вузол дерева рішень (окрім власне листя-рішень) зв'язується з деяким питанням. Будь-яке таке питання має мати кінцеве число можливих відповідей, кожна з відповідей породжує додатковий вузол на наступному рівні дерева.

– Вихід – одне з можливих рішень, яке може бути прийняте за допомогою даного алгоритму. Множина виходів збігається з множиною листя в дереві прийняття рішень. Найбільш типовим класом задач, що вирішуються методом дерева рішень, є задачі розпізнавання і класифікації. Для таких завдань дерево рішень може розглядатися як деяка ієрархія категорій, до якої можуть відноситися об'єкти, що класифікуються. Завдання методу – визначити найбільш вузьку категорію (лист), до якої із заданою ймовірністю відноситься аналізований об'єкт. В роботі [8] алгоритм дерева рішень може застосовуватися на всіх етапах аналізу тексту, використовуючи в кожному випадку різні типи передісторії, набори запитань і множини виходів. У даній роботі алгоритм дерева рішень розглядається стосовно морфологічного аналізу.

Загальна схема використання методу дерева рішень в задачі встановлення ЗКЗ виглядає таким чином:

1. Визначається множина об'єктів, їх морфологічні характеристики і множина умов, доступних для перевірки у рамках вирішуваної задачі.

2. На основі одержаної множини задіюється алгоритм побудови дерева рішень, в ході якого визначається набір питань, форма дерева і вагові коефіцієнти для кожного вузла.

3. Побудоване дерево застосовується до нового прикладу передісторії, для якого вимагається знайти рішення.

Опис алгоритму побудови дерева рішень

Як вхідні дані для алгоритму ДПР використано множину квантів S розміру S_r, множину слів-кандидатів на зіставлення C розміру C_r і множину булевих питань q_i. Булеві питання q_i являють собою тести на визначення тотожності між елементами множини C. На виході алгоритму – ймовірнісні коефіцієнти змістової залежності між елементами множини S.

Формально алгоритм побудови дерева рішень можна описати так.

Для вершини $k_{i\in[1;Sr]} \in S$ визначити елементи, які належать множині C: $x_{i1},...,x_{ij} \in C$.

Оцінити імовірності змістової залежності між елементами множини S за формулою (1):

$$P_{zk}(k_i, k_{m_t}) = \{1\text{-}\alpha * (t\text{-}1) : P(x_{ij} \equiv x_{m_t 1}) \to 1\}, \quad (1)$$

де $m_t \in [i+1; C_r]$ – ідентифікатори КНІ, які змістовно-зв’язні з k_i КНІ;

$m_{i+1} = \{\min f : P(x_{ij} \equiv x_{fl}) \to 1,\ f \in [i+1; C_r]\}$ – ідентифікатор КНІ, для якого ймовірність змістової залежності від k_i-го КНІ є найвищою;

$\alpha \in (0;1]$ – крок спадання ймовірності змістової залежності;

$t = 1,...,[\frac{1}{\alpha}]$ – кількість змістовно-зв’язних КНІ.

Ключовим моментом при використанні алгоритму побудови дерева рішень є вибір критерію завершення. Критерій завершення – це деяке правило, що визначає, в який момент алгоритм повинен припинити подальше «розщеплення» вершин дерева. Для розробленого морфологічного аналізатора критерієм завершеності формування вершини є:

1. $P_{zk}(k_i, k_{m_t}) \underset{t=2,..,\frac{1}{\alpha}}{\to} 0$, якщо для вершини k_i $\exists m_1 : P_{zk}(k_i, k_{m_1}) \to 1$;

2. $m_t = C_r$ (повний перебір елементів множини C) при невиконанні першої умови.

Лише після цього вершина вважається остаточно сформованою, і алгоритм переходить до формування наступної вершини.

Алгоритм формування дерева рішень:

– стартовою визначити вершину k_1 – корінь дерева.

– поки для даної вершини не спрацював критерій завершеності:

1. визначити дочірню вершину k_2;

2. для вершин k_1 і k_2 застосувати правило (1) зі значеннями $k_i = k_1, k_{m_t} = k_2$;

3. рекурсивно повторити алгоритм, послідовно змінюючи дочірню вершину.

Перевірка здійснюється відповідно до індексів КНІ.

У результаті роботи морфологічного аналізатора з кожним КНІ зв’язується дві множини:

1. перелік КНІ і відповідні ЗКЗ, які змістовно залежать від даного КНІ – схема «зверху вниз», дана схема формує свого роду доповнення даного КНІ;

2. перелік КНІ і відповідні ЗКЗ, він яких даний КНІ змістовно залежить – схема «знизу вверх»; дана схема породжує необхідну для засвоєння даного КНІ множину КНІ.

Встановлені ЗКЗ використовуються у двох напрямах:

1. від ключових понять дисципліни (чи окремого заняття) до необхідних для засвоєння НБ;

2. від встановлених з допомогою тестів недостатньо засвоєних НБ до НБ, від яких вони змістовно залежать.

В обох випадках застосовуються дві наведені вище схеми.

Ключові поняття (КП) – це поняття, достатнє засвоєння яких є необхідним для зарахування дисципліни. Ці КП є свого роду програмовими вимогами для підсумкового контролю. Для полегшення і пришвидшення встановлення змістових зв'язків між КП і КНІ пропонованого матеріалу, для кожного КП зазначається ідентифікатор навчального заняття (чи кількох занять).

Визначення якісних КЗ

На початковому етапі якісні КЗ визначаються експертом чи експертною системою, але за необхідності їх уточнюють в процесі засвоєння студентами НМ. Формула уточнення залежності досягнення певної якості засвоєння *j*-го КНІ від якості засвоєння *i*-го КНІ має вигляд:

$$jk_{i,j} = \frac{\sum_{m=1}^{l} r_{i,j,m}}{l},$$

де
$$r_{i,j,m} = \begin{cases} \dfrac{z_{i,m}}{z_{j,m}}, \text{якщо } z_{i,m} \le z_{j,m} \\ \dfrac{z_{j,m}}{z_{i,m}}, \text{якщо } z_{j,m} < z_{i,m} \end{cases}$$

l – кількість одночасних вивчень і-го та j-го КНІ;

m – порядковий номер одночасного вивчення *i*-го та *j*-го КНІ $m \in [1; l]$;

$z_{i,m}, z_{j,m}$ – ступінь засвоєння *i*-го (*j*-го) КНІ при *m* вивченні, $z_{i,m}, z_{j,m} \in [0;1]$;

$r_{i,j,m}$ – якісний КЗ між *i*-м та *j*-м КНІ при *m* вивченні.

Чим більше число *m*, тим точніше визначається *jk* між ступенями засвоєння КНІ, оскільки виключаються випадкові залежності.

Значення КЗ для всіх трьох типів зв'язку фіксуються у вигляді трьох різних матриць, елементами яких є відповідні значення КЗ:

	V_1	V_2	V_3	...	V_n
V_1	1	$k_{1/2}$	$k_{1/3}$	...	$k_{1/n}$
V_2	$k_{2/1}$	1	$k_{2/3}$	...	$k_{2/n}$
V_3	$k_{3/1}$	$k_{3/2}$	1	...	$k_{3/n}$
...	...	...	...	...	...
V_n	$k_{n/1}$	$k_{n/2}$	$k_{n/3}$	...	1

де V_i – ідентифікатори КНІ, $k_{i/j}$ – значення КЗ [5].

Визначення ступеня важливості КНІ

Ступінь важливості КНІ визначається виходячи з твердження: чим більше КНІ змістовно і якісно пов'язаний з ключовими поняттями дисципліни, тим вищий його ступінь важливості(S_v) :

$$S_v = \alpha \frac{\sum_{l=1}^{n} jk_l}{n} + (1-\alpha)\frac{\sum_{k=1}^{m} zk_k}{m},$$

де jk_l – l-й якісний КЗ;

n – кількість якісних КЗ, значення яких вище деякого порогового значення j;

zk_k – k-й змістовий КЗ;

m – кількість змістових КЗ, значення яких вище деякого порогового значення zd;

$\alpha \in [0;1]$ – зважений коефіцієнт впливу якісних та змістових КЗ.

Визначення ступеня складності КНІ

Ступінь складності КНІ обчислюється виходячи з твердження: складність КНІ тим вища, чим нижча ймовірність пройти його на високий ступінь засвоєння за першою спробою.

Нехай подія А – проходження КНІ на високий ступінь засвоєння за першою спробою. Оцінка ймовірності настання події А:

$$P(A) = \frac{k_1(A)}{k},$$

де k – загальна кількість студентів, що проходили КНІ у першій спробі;

$k_1(A)$ – кількість тих, що у 1-й спробі пройшли заняття на високий ступінь засвоєння.

Ці дані беруться зі статистики проходження досліджуваного КНІ.

Побудова індивідуальної траєкторії навчання

Використання встановлених параметрів КНІ і відповідних параметрів моделі студента дає можливість побудови індивідуальної траєкторії навчання в АСДН.

Виходячи з переліку ключових понять, які є кінцевою метою проходження дисципліни і незнання яких визначено з початкового тестування, формується множина КНІ, які містять дані поняття. Ці КНІ називаються ключовими. Виходячи з КЗ, визначаються спрямовуючі (допоміжні) КНІ, які є необхідними для достатнього засвоєння ключових КНІ.

Алгоритм адаптивного формування складових ЛМ:

1. Визначаються ключові КНІ.

2. Використовуючи значення якісних КЗ, формується множина КНІ, від яких ключові КНІ якісно залежать (якісні КНІ).

3. Формується множина КНІ, від яких синтаксично залежать ключові і якісні КНІ.

4. Відповідно до встановлених зв'язків і сформованих множин КНІ будується послідовність навчання.

Це є кістяк НМ. Встановлені КЗ, ступені важливості та складності дають можливість коригувати зміст, складність і послідовність викладу НМ відповідно до моделі студента. Поступове введення змістовних, менш якісно залежних КНІ чи роз'яснень існуючих знижує складність заняття. Встановлені зв'язки дозволяють компактніше розмістити більш зв'язні КНІ, сприяючи там самим кращому їх засвоєнню. Причому КНІ, які належать одному ННБ повинні подаватися лиш в межах цього ННБ, а їх кількість в межах даного ННБ варіюється відповідно до встановлених параметрів. Послідовність навчання змінюється за рахунок зміни послідовності викладу ННБ різних рівнів ієрархії.

Висновки

Таким чином, розроблений і створений програмний модуль спрощує процедуру підготовки лекційного матеріалу до імпорту в БЗ, встановлення взаємозв'язків між окремими елементами БЗ, визначення властивостей і їх подальше використання для побудови індивідуальної траєкторії навчання. Його ефективність апробовано в процесі імпорту лекційного матеріалу в БЗ АСДН.

Розроблений морфологічний аналізатор спрямований на підвищення ефективності функціонування БЗ і раціональне використання ресурсів систем навчання. Він може бути застосований не тільки в системах дистанційного навчання, але і при здійсненні традиційного процесу навчання для допомоги вчителю чи викладачу у побудові навчальних занять.

Література

1. Ong J. An Intelligent Tutoring System Approach to Adaptive Instructional Systems // J. Ong, S. Ramachandran. – United States Army Research Institute for the Behavior and Social Sciences, 2005. – 54 с.
2. Субботін С.О. Подання й обробка знань у системах штучного інтелекту та підтримки прийняття рішень : навчальний посібник // Субботін С.О. – Запоріжжя : ЗНТУ, 2008. – 341 с.
3. Технічний опис спеціальних позначень стандарту імпорту/експорту "EduPRO/Owl" [Електронний ресурс]. – [Перше видання]. – Режим доступу : http://wenet.pu.if.ua/?mod=blog&action=ReviewOne Article&id_ba=18&page=1
4. Minsky M. A Framework for Representing Knowledge / Minsky M. – MIT-AI Laboratory Memo 306, 1974. – 76 с.
5. Федорук П.І. Модель процесу прийняття рішень при побудові індивідуальної навчальної траєкторії в адаптивних системах дистанційного навчання / П.І. Федорук, М.С. Дутчак, М.В. Пікуляк // Системи підтримки прийняття рішень. Теорія і практика : збірник доповідей науково-практичної конференції з міжнародною участю. – Київ, 2010. – С. 216-220.
6. Режим доступу : http://lcorp.ulif.org.ua/dictua/
7. Андреев А.М. Вероятностный синтаксический анализатор для информационно-поисковой системы // А.М. Андреев, Д.В. Березкин, А.В. Брик, Ю.А. Кантонистов // Компьютерная хроника. – 1999. – № 1.
8. Magerman D.M. Natural Language Parsing as Statistical Pattern Recognition [Електронний ресурс] / D.M. Magerman // A dissertation submitted to the department of computer science at the committee on graduate studies of Stanford University, 1994. – Режим доступу : www.xxx.lang.gov/cmp.lg.

Literatura

1. J. Ong. An Intelligent Tutoring System Approach to Adaptive Instructional Systems// J. Ong, S. Ramachandran. United States Army Research Institute for the Behavior and Social Sciences. 2005. 54 s.
2. Subbotin S. O. Podannya y obrobka znan u sistemah shtuchnogo intelektu ta pidtrimki priynyattya rishen: Navchalniy posibnik. Zaporizhzhya: ZNTU. 2008. 341 s.
3. Tehnichnyj opys spetsialnyh poznachen standartu importu/eksportu "EduPRO/Owl". http://wenet.pu.if.ua/?mod=blog&action=ReviewOneArticle&id_ba=18&page=1
4. Minsky M. A Framework for Representing Knowledge.MIT-AI Laboratory Memo 306. 1974.76 s.
5. Fedoruk P.I. Systemy pidtrymky pryjnyattya rishen'. Teorija i praktyka. Zbirnyk dopovidej naukovo-praktichnoji konferenciji z mizhnarodnoyu uchastju. Kyiv. 2010 r. S.216-220.
6. http://lcorp.ulif.org.ua/dictua/ (morfologIchniy onlayn slovnik)
7. Andreev A.M. Kompyuternaya hronika № 1. 1999.
8. Magerman D.M. Natural Language Parsing as Statistical Pattern Recognition. A dissertation submitted to the department of computer science at the committee on graduate studies of Stanford University. 1994 www.xxx.lang.gov/cmp.lg.

RESUME

P.I. Fedoruk, M.S. Dutchak

Intelligent Analysis of Natural-Language Texts in the Process of Building of Knowledge Base for Adaptive E-Learning

In the article, the algorithm and software implementation of the module for an automated quantization of the lecture material based on natural-language texts in the process of building of knowledge base for adaptive e-learning are described. The result of this study is the simplification of procedures for the preparation of lecture material to be imported into the knowledge base, which includes the structuring of the lecture material, establishing of the relationship between the individual elements of the knowledge base, and determination of the properties and their use for the construction of an individual trajectory of study.

The effectiveness of this design was tested during the import of the lecture material in the knowledge base for e-learning.

The developed morphological analyzer aimed at improving the efficiency of the knowledge base and resource management systems training. It can be used not only in distance education systems but also in the implementation of the traditional learning process to help the teacher or instructor in the construction of training.

Стаття надійшла до редакції 02.07.2012.